

CHAPTER I

Introduction

Combining forms (henceforth CFs) are traditionally defined as neoclassical elements of Greek or Latin origin, such as *bio-* (Gr. *βίο*, Lat. *bio* ‘life’) or *-logy* (Gr. *λογία*, Lat. *logia* ‘science’), which are added to other elements, either initially, as in *biodata* ‘one’s curriculum vitae’, or finally, as in *musicology* ‘the branch of knowledge that deals with music’,¹ with the addition of a linking element *-o-*. Because they are bound morphemes, scholars have raised the issue of whether neoclassical CFs should be regarded as cases of affixation or compounding. The prefix status of *bio-* and the suffix status of *-logy* are disconfirmed by the fact that the two can be combined to form a word (i.e. *biology*) which cannot consist of a prefix and a suffix without a root, but does consist of two CFs. Bauer et al. (2013: 441–442) therefore assume that neoclassical formations are to be treated as cases of compounding, rather than as cases of derivation. Bauer (2017: 150) also adduces an etymological explanation to the same assumption: since these forms are compounds in the languages in which they originate, they remain compounds when borrowed into English.

However, the overall picture is more complex and complicated, comprising not only neoclassical bound elements but also abbreviated forms, such as *e-* (from *electronic*) in *e-book* and *e-journal*, and secreted forms, such as *-gate* (from *Watergate*, denoting ‘an actual or alleged scandal’) in *Billygate* and *sexgate* (cf. classical vs modern CFs in Prčić 2005, 2008; Amiot & Dugas 2021: §§ 3.2.1–3.2.2). Abbreviated and secreted forms are bound because they involve some type of shortening (either back- or fore-clipping), but in the former meaning remains intact, while in the latter reinterpretation occurs. The question remains open as to whether they should be regarded as compounding or derivation. For instance, *-(a)holic* (from *alcoholic*) was formerly considered a CF (e.g. by Warren 1990; Lehrer 1998; Fradin 2000), but in the *Oxford English Dictionary*

¹ All definitions used are from the *Oxford English Dictionary* unless otherwise indicated.

(henceforth *OED*) it is now described as a suffix used ‘to form nouns denoting people who appear to be addicted to the thing, activity, etc., expressed by the first element’, as in *workaholic* ‘a person addicted to working’, *shopaholic* ‘a compulsive shopper’, etc.

Warren (1990: 112) was the first to stress that ‘combining forms are morphemes of a rather special kind’, which represent a quite heterogeneous group of elements, also differing in their origin as (1) allomorphs of model words (e.g. *astro-* in *astrodome*), (2) truncated forms of model words (e.g. *cyber-* in *cyberphobia*), or (3) parts of model words which happen to coincide with existing words, such as *-gate* above. The same CF can even differ in the way it is interpreted depending on the word to which it is attached: e.g. in *cheeseburger* ‘hamburger topped with cheese’, *-burger* is an abbreviated form, whereas in *fishburger* ‘fried patty made of fish served in a bread-bun’, it is secreted, in that a *fishburger* does not contain any beef. This heterogeneity manifests itself in terminological confusion (e.g. affixoids, pseudo-affixes, semi-affixes, affix-like formatives, folkmorphs, and splinters, along with CFs)² and descriptive vagueness, in addition to the lack of a morphological account that embraces the various types. Bauer already claimed in the 1980s that CFs belong to a type of word-formation which ‘has received very scant attention in the literature on morphology’ (Bauer 1983: 213). More recently, Kastovsky (2009: 12) has argued that the notion of CFs is not necessary at all, as ‘the categories of “word”, “stem”, “affix”, “affixoid”, “clipping” and “blending” necessary in word-formation for independent reasons are sufficient to deal with the formations in question’. As Bauer et al. (2019: 62) observe, CFs remain a ‘fuzzy’ category, but nonetheless a necessary one deserving investigation, clarification, and positioning within the morphological ecosystem of modern English.

1.1 Overview and Rationale of the Work

This work aims to fill the descriptive and theoretical lacuna surrounding CFs as well as to offer a broad spectrum along which new English CFs can be arranged. CFs are defined here as initial or final bound morphemes which are either allomorphic variants of classical Latin or Greek words

² Cf. It. *elemento formativo* ‘formative element’ in Iacobini (2004), *prefissoide* ‘prefixoid’ and *suffissoide* ‘suffixoid’ in Migliorini (1963), or even *semiparola* ‘semiword’ coined by Scalise (1983). In French, Tournier (1985) uses the term *fractomorphème* ‘fractomorpheme’, while Renner (2008) prefers using *quasi-lexème* ‘quasi-lexeme’ and Ronneberger-Sibold (2010) uses the German term *Konfixe* for ‘combining forms’. Hamans (2017) replaces the term ‘splinter’ with ‘libfix’, a neologism (from the ‘liberation’ of parts of words) which he borrows from Zwicky (2010).

Overview and Rationale of the Work

3

(e.g. *bio-*, *-logy*), or shortenings of (native or non-native) English words (e.g. *e-* from *electronic*, *econo-* from *economic/economy*), often with the intervention of a secretion process (e.g. *-(a)holic* ‘person addicted to’, *-gate* ‘political scandal’). These are respectively called ‘neoclassical’, ‘abbreviated’, and ‘secreted CFs’, and the complex words obtained from them are called ‘combining-form combinations’ (inspired by Quirk et al.’s 1985: 1575 ‘combining-form compounds’), including neoclassical compounds (e.g. *bio-therapy*, *zoo-logy*), as well as combinations with abbreviated initial (e.g. *e-reader*, *econo-politics*) or final secreted (e.g. *computer-holic*, *oil-gate*) elements.

In this work, the three categories of neoclassical, abbreviated, and secreted CFs are viewed as part of transitional morphology, i.e. intermediate between two subcomponents of word-formation: compounding and derivation. The label ‘transitional morphology’ is preferred to ‘marginal morphology’, used by Dressler (2000: 1) to refer to phenomena that, within the naturalness theory, lie at the boundaries of grammatical morphology. The notion of marginality indeed not only presupposes the presence of (discrete or fluid) boundaries but also conveys a negative connotation of being ‘less important’, ‘trivial’, or ‘secondary’, for example compared to grammatical word-formation, which is not the case with CFs, which are grammatical and productive in English. In transitional morphology, instead, there are transitions between the external and internal boundaries of morphology: e.g. between morphology and other components, such as phonology, syntax, semantics, pragmatics, or between sub-components of word-formation, such as composition and derivation. Transitions, therefore, imply the notions of (1) dynamism and directionality, (2) boundaries, (3) prototypicality vs non-prototypicality, and (4) graduality vs dualism/superposition (see § 2.2.4).

There is considerable variability among the aforementioned categories of CFs in terms of prototypicality, productivity, and profitability (Plag 1999; Bauer 2001). In this work, more established (classical) and newly created (modern) CFs will be kept distinct, but I will verify the hypothesis that they represent a continuum rather than separate classes of word-formation, as some of them are non-prototypical representatives of compounding, while others are non-prototypical representatives of affixation.

A fourth category which will be discussed in the work consists of potential CFs, i.e. blend splinters. Indeed, the shortening process at the origin of abbreviated and secreted CFs is often connected with blending. Blends can be defined as lexical items intentionally formed by merging word parts (called ‘splinters’, Lehrer 1996, 2007), usually from two lexical source units

generally known as ‘source words’ (e.g. *jeggings* ← *j*(eans) + (*l*)eggings) (Mattiello 2021: 6). However, when blend splinters are used recurrently and become productive, they can start their life as independent morphemes. As argued in Mattiello (2017), splinters such as *docu-* (from *documentary*) in *docudrama* ‘a documentary drama’ or *-cation* (from *vacation*) in *staycation* ‘a holiday spent at home’ may give rise to analogical formations or productive series. For instance, the initial splinter *docu-* may also be found in other blends such as *docusoap* and *docutainment* (*OED*), and similarly the final splinter *-cation* is attested in *workcation* ‘a vacation in which you bring your work with you’ and *spa-cation* ‘a spa vacation’ (*Urban Dictionary*). As substantiated by contemporary corpus data in Bauer et al. (2019) and Beliaeva (2019), the morphological status of formations containing splinters can be compared to that of neoclassical compounds, or even to affixations. Moreover, their creation of series sharing the same initial or final element can confirm their productivity. Hence, in this work blend splinters will be treated as CFs in the making (Chapter 7): ‘in the making’ because they have not yet been included in the *OED* as separate entries, although their frequency in novel blends and their regularity can motivate their status as ‘proper CFs’. Unlike word parts in blends, however, secreted forms acquire a level of abstraction (through semantic generalisation or specification) which is not typical of blends, and which contributes to their transition towards affixation. In other words, blend splinters are also part of a continuum from a word part used in the coinage of a blend to a frequent splinter used in series up to a recognised abbreviated (e.g. *docu-* or *-cation*) or even secreted (e.g. *-ercise* or *-(t)arian*) CF.

For this reason, in recent times the term ‘splinter’ has shifted from its original sense of any portion of a base word that has become part of a blend (Adams 1973; Bertinetto 2001; López Rúa 2004; Gries 2012) to a portion of a word which has been attached recurrently to new bases (Bauer 2006; Bauer et al. 2013; Beliaeva 2019) or to ‘a new type of bound morpheme that results from secretion’ (Callius 2016: 513), as is the case with *docu-* or *-cation*. Nevertheless, some combinations with splinters are closer to blends – e.g. *docu-* is attached to another splinter in *docutainment* (i.e. *-tainment* from *entertainment*) – while others are added to free words (*work*, *spa*), as typically happens within productive and regular derivation. Therefore, we can assume that splinters represent different points on a cline between creativity and productivity (see Dal & Namer 2017 below for ‘creativity’). As observed by Correia Saavedra (2013: 12), ‘what seems to be a relevant factor in the transition from a blend splinter to a combining form, is the notion of repetition and frequency’. This position

is also supported by Lehrer (2003) and Renner (2008), as well as by Fischer (1998), who even quantifies frequency: ‘At the bare minimum, at least three neologisms must be found containing the blend element before the existence of a new combining form can be assumed’ (p. 65). Put in these terms, splinters can be considered CFs if they display frequency and productivity.

As observed by Lehrer (1998: 4, 2007: 121), once a splinter has demonstrated some productivity, it can result in ‘a morpheme in its own right’, that is, in ‘a new combining form’. This claim broadens the view on CFs, which depart from the original taxonomies of neoclassical forms (Marchand 1969; Bauer 1983; Plag 2003) and widen to embrace recently created new forms (Lehrer 2007; Amiot & Dugas 2021).

The rationale behind this work is to provide a theoretical model for CFs in English, clearly distinguishing them from apparently neighbouring categories, such as abbreviations or blending, which instead belong to ‘extra-grammatical morphology’ (Mattiello 2013), or derivation and compounding, which are rather part of prototypical ‘morphological grammar’ (Dressler 2000: 1). As an additional goal, this work aims to provide an overview of new CF combinations that are expanding modern English vocabulary.

With these goals in mind, the present work will address the following research questions:

1. Are CFs a necessary and independent morphological category and, if so, what is their locus (a) in modern English word-formation theory in general and (b) in relation to derivation and compounding in particular?
2. Are CFs a heterogeneous group of elements, or do they rather belong to different subtypes with their own specific features, frequencies, and productivity?
3. What is the profitability of CFs in modern English? Are they enriching the lexicon with new established words, with nonce words, or both?
4. Are there novel splinters that are potential candidates for the role of CFs?

When new words are coined, it is often a matter of creativity in that novel items entering the lexicon originate from the creative mind of the speakers. According to Dal and Namer (2017: 223), the distinction between productivity and creativity ‘is connected with the disputed opposition between intentionality and unintentionality in word-formation’. They reserve the term ‘creativity’ for cases in which the coined word

‘obviously transgresses the morphological system, such as in poetry or playful creations’ (Dal & Namer 2017: 223), though admitting that playful coinages can also use productive rules.

This book shows that CFs are not the result of ‘extra-grammatical word creation’ in Ronneberger-Sibold’s (2010: 201; 2015) terms, but rather belong to that part of morphology displaying regularity, predictability, and productivity, as they regularly contribute to making new forms available and profitable (see ‘profitable’ vs ‘available’ in Carstairs-McCarthy 1992: 37; ‘rentable’ vs ‘disponible’ in Corbin 1987: 177). Some CF combinations are indeed mere temporary or transitory ‘nonce formations’, in Bauer’s (1983: 45) sense of ‘new complex word[s] coined by a speaker/writer on the spur of the moment to cover some immediate need’, but most of them are established neologisms which have entered the English language, developed its speakers’ vocabulary, and potentially become part of the linguistic heritage of English-speaking countries. As remarked by Miller (2014): ‘Combining forms are morphologically and semantically motivated. Their diffusion via existing words with the same element leads to productivity and promotes institutionalization’ (p. 213).

1.2 Organisation of the Work

This work is organised into six main chapters. Chapter 2 explores the literature on CFs and related phenomena, from neoclassical compounds to splinters. This chapter also deals with notions that are crucial in the classification of CFs, such as the bipartition between extra-grammatical and grammatical morphology, and the further subdivision of the latter into prototypical and central on the one hand, vs non-prototypical and transitional on the other. It also shows the relationship between transitional morphology and paradigmatic as well as extravagant morphology, and finally operationalises the concept of ‘transitional morphology’, with reference to various languages but a focus on the phenomena of transitional morphology in English. The chapter offers a working model for the categorisation of CFs, where these are distinguished from affixes, affixoids, and compounds, but partially overlap with them, displaying some of their characteristics. The part of morphology to which CFs belong is called ‘transitional morphology’; in particular, they are at the borders between derivation and compounding. The identification of a dividing line between abbreviated and secreted CFs is another issue that this chapter addresses, by introducing the concepts of morphological productivity, abbreviation vs secretion, analogy and schema, reanalysis, and semantic weight.

Chapter 3 provides an overview of the dataset collected for the linguistic analysis and an explanation of the methodology used for data selection, as well as an enumeration of the corpora investigated for quantitative and qualitative study. The chapter also introduces the corpus-based approach to morphological productivity adopted for the quantitative investigation.

Each of the following three chapters is devoted to one of the different categories of CFs taken into consideration – i.e. neoclassical (Chapter 4), abbreviated (Chapter 5), and secreted (Chapter 6) – alternating descriptive accounts with corpus-based analyses which attest the actual and current use of these forms. Chapter 7 finally focuses on more recent forms still fluctuating between the status of blend splinter and that of abbreviated or secreted CF. This chapter shows that the transition from a blend splinter to an abbreviated CF can be determined by the recurrence of the splinter in series of novel words. When the CF acquires abstraction and either generalisation or specification, it eventually becomes secreted.

Quantitative analyses can shed light on the role that these forms play in language and in the morphological ecosystem. Each chapter also provides instances of CF combinations in genuine contexts and different textual genres, ranging from academic discourse to newspapers and magazines, and from journals to blogs and other web texts.

1.3 Readership of the Book

Let us conclude this introductory chapter with some marginalia on the target readership of this book. The book addresses specialists in linguistics and non-specialists alike. First, it addresses an audience of morphologists and experts in word-formation, as well as scholars and learners of English, in that CFs have been disregarded for too long hitherto and a theoretical model embracing them is needed in the description of modern English grammar. Second, it addresses lexicographers and lexicologists, who may face the challenging task of classifying new words as either derived or compound complex forms, or introducing bound morphemes as new entries in their dictionaries. Results from corpus-based analyses may encourage exclusion or inclusion, and in the latter case, may help lexicographers to label bound morphemes as either ‘suffixes’, ‘combining forms’, or rather ‘blend splinters’. Moreover, this book addresses experts other than linguists: for instance, many new neoclassical CFs are nowadays part of scientific terminology and thus might be of interest to researchers in the medical field, doctors, or scientists. Still another portion of the public may be newspaper or magazine journalists interested in attracting their readers’

attention by means of novel and often amusing vocabulary. Think for instance of all the neologisms and nonce words that have been coined by prefixing *Brit-* (from *British*) to nouns (e.g. *Britart*, *Britcom*, *Britculture*, *Britfilm*, *Britflick*, *Britlit*, *Britpop*), or the wordplays created in the news after the *Brexit* phenomenon (e.g. *Frexit*, *Italexit*, and *Spexit*, respectively referring to ‘French, Italian, or Spanish withdrawal from the EU’) (Mattiello 2019a). Are *Brit-* and *-exit* novel CFs? What is their role in the lexical expansion of English vocabulary? Are they attested in dictionaries and, if not, why not? The theoretical model adopted here will guide readers beyond the confines of the scholarly community in addressing (and hopefully answering) these and related questions on English CFs.

CHAPTER 2

Background of Combining Forms

CFs are notoriously difficult to define and classify, and taxonomies generally focus on one or the other type without providing an all-inclusive framework. A distinction is often made in the literature between initial CFs (henceforth ICFs) and final CFs (henceforth FCFs) (Bauer 1983; Warren 1990; Plag 2003), depending on their position. Another commonly shared conception among scholars is that CFs are bound morphemes, which as a result connects them to other bound forms. Three commonly recognised types of bound morphemes are affixes (e.g. *-ic* in *alcoholic*, *re-* in *rewrite*), roots (*rac-* in *racism*, *prob-* in *probable*), and neoclassical CFs such as *astro-* or *geo-* initially and *-cide* or *-itis* finally, which are exemplified by neoclassical compounds (*astro-logy*, *geo-graphic*; *geno-cide*, *laryng-itis*). However, the question of how to classify and characterise bound morphemes like *porta-* (from *portable*, in *porta-kit*, *porta-office*) and *-tainment* (from *entertainment*, in *edutainment*, *infotainment*) is challenging. They may occur either initially or finally, but certainly they are not the same as the root *prob-* or the suffix *-ic* above, nor do they exactly correspond to neoclassical forms. Moreover, *porta-* attaches to words (*kit*, *office*) and *-tainment* is added to splinters (*edu-* from *education*, *info-* from *information*), whereas *astro-* or *geno-* can combine with other neoclassical CFs, i.e. *-logy* and *-cide*.

In this chapter, the literature on CFs is inspected in order to shed light on the distinction between CFs and other adjacent bound and unbound forms, as well as to provide a model for their classification and description.

2.1 Survey of Definitions and Descriptions of Combining Forms

The literature on English morphology offers narrower and broader accounts of CFs and CF combinations, which for the most part do not provide a clear-cut distinction among the three categories taken into consideration here, but mainly focus on one or the other.

2.1.1 *Marchand (1969) and Adams (1973)*

While discussing prefixing on a neo-Latin basis of coinage, Marchand (1969: 131) uses the term ‘combining forms’ to refer to particles that are ‘stems of full words in Latin or Greek’, such as *multi-* (Lat. *multus* ‘much, many’), *omni-* (Lat. *omnis* ‘all’), *astro-* (Gr. *ἄστρον* ‘star’), and *hydro-* (Gr. *ὑδωρ* ‘water’). These represent loans in English, with no independent existence as words. However, the fact that they were independent in their source languages makes us assume that the lexemes *multi-word*, *omni-science*, *astronomy*, and *hydrospace* can be considered compounds rather than prefixed words. On the other hand, Marchand (1969: 132) observes that the lexicographic distinction between the CFs *multi-* or *hydro-* and prefixes such as *hyper-* or *hypo-* is ‘by no means justified’. The latter, indeed, have a Greek origin, and combine with both free words (e.g. *hyper-active*, *hypo-thesis*) and neoclassical CFs (e.g. *hyper-trophy*, *hypo-drome*), which is not possible for prefixes according to Bauer (1983: 214–215). As a solution to this thorny problem, Warren (1990: 113) proposes to consider *hyper-* a prefix on some occasions (e.g. *hyperactive*), but a CF on others (e.g. *hypertrophy*). However, this would admit that morphemes like *hyper-* have an ambiguous nature, without explaining the distinction between affixes and CFs (see § 2.4.1).¹

As for suffixing, Marchand (1969: 356–358) does not consider final forms of neo-Latin basis, such as *-graphy* or *-logy*, but only discusses in his account what he terms ‘semi-suffixes’, i.e. elements that ‘stand midway between full words and suffixes’, including *-like*, *-worthy*, *-monger*, *-way(s)*, and *-wise*. However, in the *OED* these do not represent a homogeneous group. For instance, *-worthy* (in *noteworthy*) and *-monger* (in *fashionmonger*) are considered ‘compound constituents’; *-like* (in *manlike*) is labelled a ‘suffix’; and *-way* (in *sideway*) and *-wise* (in *foodwise*) are defined as ‘combining forms’. This terminological confusion is accompanied by the fact that all of these elements are also used as autonomous words, suggesting that combinations with them are not suffixations, as Marchand’s term ‘semi-suffix’ would imply, but rather compounds. As argued by Dalton-Puffer and Plag (2000: 243), ‘there are good arguments for treating these formations as either compounds (in the case of *-type*), or as suffixes (in the cases of *-ful* and *-wise*). Labels such as “semi-suffix” are theoretically undesirable and do not provide additional insights into the nature of complex words.’

¹ Cf. Amiot and Dugas (2021: § 3.1), who rule out *hyper-* from their discussion on CFs because it does not correspond to a lexeme but to a grammatical word.